Prediction of the Compression Index of Sudanese Soils Using the Neural Network Systems

A Thesis Submitted to the Department of Civil Engineering, Faculty of Engineering and Architecture, University of Khartoum in Partial Fulfillment to the Requirements of the Degree of Master of Science in Structural Engineering.

By
Ahmed Hassan Mohamed Ahmed (B. Sc. (Eng,))
Nov. 2003

Dedication

To my dear parents and caring wife with love..

..And to all those who reside my heart

..and whose hearts I reside.

Contents

Acknowledgement

The author feels deeply gratified to many persons he had encountered on his journey with this research bit without whom, save for Allah, this thesis might never have seen the light. So, the utmost gratitude and hearty thanks are due to Dr. Ahmed Mohamed El Sharif, the Supervisor, for his kind guidance, patient follow-up and precious comments. More thanks are due to the staff of the BRRI library and on top of them Mr. Mobarak A/ Atti for his gentle and helpful cooperation. Due thanks and precious gratefulness to Eng. Asim A/ Sansoi for his well known, yet highly appreciated- helpfulness.

And the thanks are paid to all those who are- unintentionally- not mentioned with all hopes that Allah may bless them.

Abstract

This thesis is concerned with evaluation of the compression index using the neural network technique. The thesis begins with browsing the researches carried to obtain a solution that suits as many as possible soil types. These efforts, in general, are concentrated on the most popular parameters that affect the compression index. The basic concepts of the neural network systems are reviewed. The results of the neural network are compared to regression analysis applied to the same data. All the data used in the thesis is collected from the service reports of the Building and Road Researches Institute, University of Khartoum.

The values of the compression index obtained from the neural network came in a close agreement to those derived from laboratory works.

List of Notations

All the notations are defined at their first appearance in the text. Here are some notations that are frequently used:

C_c=the compression index;

C_r=the recompression index;

e_o=the initial void ratio;

e= the instantaneous void ratio;

G_s=the specific gravity of the soil;

LL= the liquid limit of the soil;

MC=the initial moisture content;

PI=the plasticity index;

W_n, W=the natural moisture content;

γ_d=the soil dry density;

γ_b=the soil bulk density;

γ_w=the unit weight of water;

σ_{vo}, σ_o=the overburden pressure;

σ_p=the preconsolidation pressure;

List of figures

List of Tables

Chapter One

Introduction

Chapter (1)
Introduction

1.1 Statement of the problem

The problem of dealing with engineering materials is an ever-lasting one. The mathematical approaches were always the trusted tool to overcome the problem. These approaches involve a lot of assumptions and approximations to bring the actual state close to the theoretical one. There are many difficulties that encounter data optimization and mathematical formulations. One of these difficulties is the huge amount of calculations needed, which is time consuming and sophisticated.

The evaluation of the compression index of the soil, a measure of the compressibility of the soil, is one of the victims of the mentioned difficulties

The use of computer has made huge advances in solving the problem. Computer- based techniques are common nowadays, one of which is the artificial neural network. This thesis presents an application of the neural network systems to predict the compression index of the soil.

1.2 Objective and methodology

It is intended in this thesis to evaluate the compression index by construction of a neural network system. The network used is a backpropagation network. A regression analysis is also introduced to the compression index and its different parameters to be compared to the net results

1.3 Layout of the thesis

The thesis falls into six chapters and one appendix: Chapter One is an introduction that discusses the statement of the problem, the objective and the layout of the thesis. Chapter Two covers a literature review about the compression index. In Chapter Three, the basic concepts of the neural networks are viewed. The data collection and methodology are presented in Chapter Four. Chapter Five contains the results of both regression analysis and the constructed neural network. Finally, the summary conclusion, and recommendations are included in Chapter Six.

Chapter Two

Literature Review

Chapter (2)

Literature Review

2.1 Introduction

Soils are considered to be elastic materials, but this theory is challenged by the fact that the interrelation between stress, strain and time are not simple and cannot be treated mathematically (Holtz and Kovacs,1981), besides soils are non- conservative materials. So, the deformations that take place in the soil are permanent.

These deformations occur as a change in shape or volume or a combination of both. The sources of deformations are the loads applied to the soil from structures, fills, excavations or the fall of water table (Craig,1987). And in fine- grained soils time has its considerable effect (Terzagi, Peck and Mesri, 1996).

Generally, most of these deformations are vertical and they are called settlement. Settlement is estimated by assuming that the soil has no lateral deformation (confined soil); i.e. a one- dimensional deformation.

The one- dimensional deformation of soil results from:

1- deformation of grains.
2- compression of air and water in the voids, and/ or
3- squeezing out of water or air from the voids.

The deformation of the soil grains is very small and is neglected; hence, the one- dimensional compression of soil is caused by the change in the void volume.

The settlement is affected greatly by the permeability and the structure of the grains of the soil. For low- permeability soils (e.g. clays) under loading, the rate of squeezing pore water is the controlling factor of compression; a process called consolidation (Terzagi, Peck and Mesri, 1996).

2.2 One- dimensional compression

This term is confined mostly in saturated silts and clays. As mentioned before, the soil structure is a controlling factor of settlement. Soil structure can be identified by soil composition and effective stress history. The existence of interparticle bonds makes the soil structure response to the external effective stress, σ_v, time dependent. The stress at the existing state is the effective overburden pressure, σ'_{vo}, and the void ratio is e_o (Terzagi, Peck, and Mesri, 1996).

Soils with high volume of voids have high potential for volume changes, which decreases when the soil structure rearranges under increasing effective stress. The stress that makes the major changes in the structure of the soil is called the preconsolidation pressure, σ'_p. The range between σ'_{vo} and σ'_p is the recompression range in which no interparticle displacements result. In the range beyond σ'_p, called the compression range, the soil particles rearrange to resist the additional effective stress, so additional compression will develop to accommodate the additional stress and the bond resistant between particles (Terzgi, Peck and Mesri,1996).

2.3 The oedometer test

It is a test that is used to determine the soil characteristics during one-dimensional consolidation or swelling. The apparatus consists of a metal ring and two porous stones that confine the soil specimen from above and below. The loads are applied through a loading cap fixed to the upper porous stone. The assembly is set in an open cell of water to which the pore water has a free access. The metal ring confines the specimen, thus imposes a condition of zero lateral strain.

At the start of the test an initial pressure, which depends on the type of the soil, is applied to the specimen. Then, the pressure is doubled after 24 hours, and the compression reading is taken off a dial gauge attached to the loading cap and also the void ratio before incrementing the load. The process is repeated for several values of pressure (each value is double the previous one). When the excess pore pressure is dissipated the corresponding pressure on the soil is its effective stress. The result of the test is presented as a plot of the void ratio against the effective stress or its logarithm. The expansion of the specimen can be measured by removing the final pressure and successive decreases in the applied pressure (Craig, 1987).

2.4 The compression index C_c

The compression index, C_c, is the slope of the linear portion of the e-log σ' plot, where e is the void ratio and σ' is the effective stress. It is one of the tools used to determine compressibility of the soil.

Likewise, the recompression index C_r is the slope of the recompression portion of e-log σ' plot.

Determination of C_c value is a difficult task, though there are few equations that are obtained to relate the compression index with the soil index properties. The problem with these equations is that they are suitable for specific areas because the soil properties change with mineralogy, effective stress history and geographical location.

Some of the efforts made to predict C_c are investigated hereafter. These efforts are reported by Herrero (1980)

a) Skempton (1944) equation

The tests are conducted on different types of clay specimens with their initial void ratio at the liquid limit yielding the following relationship:

$$C_c = 0.007\ (LL - 10) \qquad (2.1)$$

C_c = the slope of virgin compression portion of consolidation

LL= the liquid limit (%).

b) Terzagi and Peck (1948) equation

The Skempton equation is modified for use with normally loaded clay of medium to low sensitivity with the result that compression index for normally loaded clays is approximately 1.3 times the remoulded value. Hence,

$$C_c = 0.009\ (LL - 10) \qquad (2.2)$$

C_c is the compression index for undisturbed normally loaded clay.

c) Nishida (1956) equation

Based on void ratio considerations, the equation is

$$C_c = 1.15\ (e - e_o) \qquad (2.3)$$

e_o=the void ratio before pressure application.

e=the void ratio where the compression index is measured .

Assuming e_o=0.35 (for uniform rigid sphere at the closest packing), Nishida equation becomes

$$C_c = 1.15\ (e - 0.35) \qquad (2.4)$$

And in terms of natural void ratio (e_n),

$$C_c = 0.54\ (e_n - 0.35) \qquad (2.5)$$

d) Hough (1957) equation

The virgin curves of different types of soils indicate the variation of compressibility with initial void ratio. Recognizing this, Hough (1957) has performed tests on remolded specimens to find that the relation between the compression index and the initial void ratio is linear. The expression is

$$C_c = a\ (e_o - b) \qquad (2.6)$$

The factor “a” represents the slope and depends on particle shape, size and gradation.

“b” is the approximate minimum void ratio at normal conditions.

According to Hough, the thickness of soil layer is directly proportional to the ratio $C_c/(1+e_o)$ rather than C_c alone at a given stress ratio. Fadum (1948) considered the ratio $C_c/(1+e_o)$ and developed the following expression for Boston blue clay:

$$\frac{C_c}{1+e_o} = a'(w - b') \qquad (2.7)$$

w= natural moisture content (%)

A graphical correlation between the ratio $C_c/(1+e_o)$ and the natural moisture content was performed by Arango in reference to Lambe and Whitman (1969) work. They concluded that any relationship between Atterberg limits and the compression index is only an approximation.

According to Nishida, some soils disobey the relationship because the coefficients have to be determined for each soil.

e) Azzouz et al (1976)

Azzouz et al (1976) reported the following two equations

$$C_c=0.37(e_o+0.003LL-0.34) \quad (2.8)$$

$$C_c=0.009w_n+0.002LL-0.1 \quad (2.9)$$

With correlation coefficients of 0.86 and 0.81 and standard errors of 0.074 and 0.085, respectively.

f) Herrero universal C_c equation

The equation that is formulated by Herrero (1980) takes the following form:

$$C'_c = f\,(G_s\gamma_w /\ \gamma_d) \qquad (2.10)$$

The compression index pertains to the slope of the virgin portion of the laboratory consolidation curve and is described by the notation C'_c.

$$C'_c = a\ G_s\,(\gamma^2{}_w /\ \gamma^2{}_d)^b \qquad (2.11)$$

$b=1.191 \approx 1.2$ and $a= 0.141$

$$C'_c = 0.141 G_s\,(\gamma^2{}_w /\ \gamma^2{}_d)^{6/5}$$

thus, $$C'_c = 0.141 G_s\,(\gamma_w /\ \gamma_d)^{12/5} \qquad (2.12)$$

the coefficient of determination, $R^2=0.85$

The slope of the zero- air- void curves is computed using e_o instead of e_v (e_v is instantaneous void ratio) because the former can be measured by experiment while the later can only be estimated .

He also developed a model that correlates the void ratio, liquid limit and compression index which took the following form:

$$C_c = -0.156 + 0.411e_o + 0.00058LL \tag{2.13}$$

Chapter Three

Neural Networks Basic Concepts

Chapter (3)

Neural Networks Basic Concepts

3.1 Historical background

The Artificial Neural System (ANS) is an application of the neurological concept to science and engineering, in order to give solutions for certain problems(e.g. pattern recognition and diagnosis of problems from their symptoms).

The first example of the ANS occurred in the late 1950's in the form of a work done by Frank Rosenblatt on a device called the perceptron (Freeman and Skapura, 1991). In the period from 1969 until the early 1980's, there was little research work in the field. A new rise took place by David Rumelhart and James Mcclelland (1986) (Freeman and Skapura, 1991) when they published their research findings; Parallel Distributed Processing (PDP), though there was other papers that were published in the same period. The findings gave the basic knowledge needed in the field of ANS.

3.2 Neurophysiology and ANS

The neurophysiology was the source of inspiration that led to artificial neural systems. As an approach to the basic idea, it is suitable to provide some information about neurophysiology by considering a single nerve cell in the nervous system (fig.3.1).

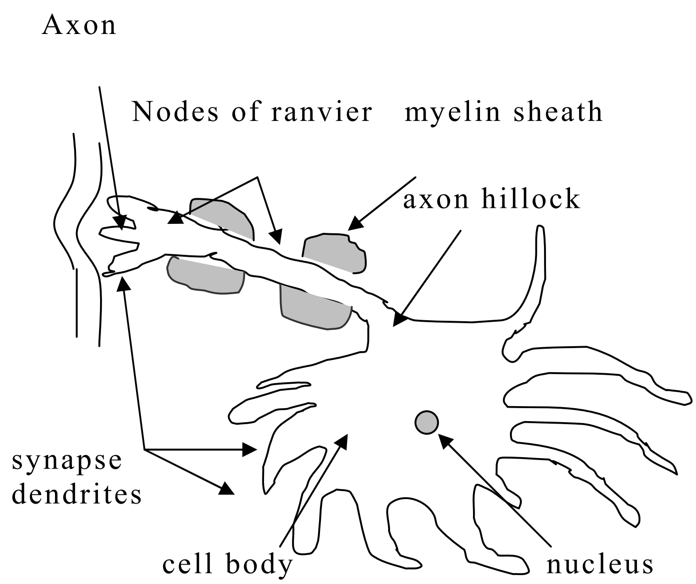


FIG. (3.1): A nerve cell

The major components of the nerve cell are:

1- the dendrites.
2- the cell body.
3- the axon.

There is a resting potential that is maintained by the cell membrane, which is permeable to certain ionic species. The membrane maintains this potential between intracellular and extra-cellular fluid by the action of a sodium- potassium pump.

When the cell is excited, a reduction in the potential difference across the membrane takes place and an action potential occurs in the cell. This action is then transmitted through the axon to other cells and the transmitting cell regains its resting potential. The neurons are connected to each other by a synapse, which passes the transmitted substance from the presynaptic cell to the synaptic cell. If the influx entering the cell tends to depolarize the resting potential, it is an excitatory otherwise it is inhibitory (Freeman and Skapura, 1991).

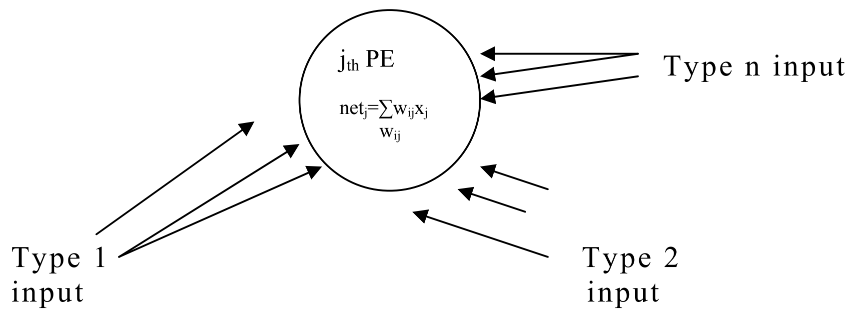
The neurons form circuits with each other in the nervous system. Each neuron sends impulses to others and receives from others, and this idea forms the basis for most neural network models. The feed back paths between neurons make the control over the system much easier, since they may be positive or negative due to the nature of the synaptic connections (Freeman and Skapura, 1991).

The learning process is a result of excitement exchanged between cells. It is believed that the changes of the excitements depend on the synaptic junction between neurons.

3.3 Artificial Neural Systems (ANS) and neurons

The individual elements in ANS are called nodes, units or processing elements (PE). And these elements represent an activity of a group (layer) of selected neurons(fig. 3.2).

Output



x_{ij} is the input to i^{th} PE from the j^{th} PE

w_{ij} is the weight on the connection between the i^{th} and j^{th} nodes

FIG. (3.2): A PE in a network

Each PE has a net input value based on all its input connections, which is expressed as

$$net_i = \sum_j x_j w_{ij} \tag{3.1}$$

j is the index for all connections to the PE.

x_j is the vector representing the output of the previous layer.

The calculated input is then converted to activation expressed as

$$a(t) = F_i(a_i(t-1), net_i(i))$$

t is an index to denote the current activation

Often the activation and the net input are identical, so the terms thereof are interchangeable

x_i is the output function=$f_i(a_i)$ or $f_i(net_i)$

3.4 ANS simulation

Network models often are built on:

1-An implementation of a unique learning law

2-An interconnected scheme, and

3-A structure

3.4.1 Types of learning

3.4.1.1 Supervised learning

The net is directed to give a certain output by applying a predefined input to it. The weights are adjusted to minimize the difference between the desired and the actual output for each input pattern (Taylor, 1995).

3.4.1.2 Reinforced learning

The net receives a global reward or penalty signal. The weights are changed in order to develop an input/ output behavior, which favors the probability of receiving a reward signal rather than a penalty signal (Taylor, 1995).

3.4.1.3 Unsupervised learning

The net discovers the statistical regularities in the input space and automatically develops different modes of behavior to represent different classes of inputs (in practice, some labeling is required after training to associate each mode of behavior with its corresponding class of input) (Taylor, 1995).

3.4.2 The Design of simulator

The most two important points to be born in mind are:

1. The software used for simulation should be such that the neural network can be sized dynamically.

2. The CPU of the computer will be busy computing the input activation value, which is a time consuming task.

The main line on the simulation design is that the units in one layer have to be fully connected to those in another layer forming a uniform interconnection (Freeman and Skapura, 1991).

3.4.3 ANS data structures

3.4.3.1 Array-based structures

The data in ANS is processed as a sum of products, thus the network data is to be arranged in groups of linearly sequential arrays containing homogeneous data. This arrangement is economical because of the easiness in dealing with arrays rather than single values, besides it makes an efficient use of memory.

The computation of aggregate input at a unit is processed by setting two pointers to the first location of the outputs and weights arrays and setting a local accumulator to zero.

The values located in memory at each of the two pointers are, then, multiplied and added to the local accumulator, incrementing both pointers, and this is repeated for all arrays' value (Freeman and Skapura, 1991).

3.4.3.2 Linked-list structures

Here the dynamic memory is implemented as lists of records containing different types of data. Each record has a pointer to the next in the chain. The list is a set of records with pointers to other records. These linked-lists provide a generality of algorithm that allow for a neural network simulation better than array structure. But it consumes more memory and run time than the array structure (Freeman and Skapura, 1991).

3.5 Back propagation network

The Back Propagation Network (BPN) is formalized by Werbos originally (Freeman and Skapura, 1991). It was found very useful in tackling problems that involve complex data pattern and nontrivial mapping function. It works as a multi-layer and feed-forward network using the supervised mode of learning.

3.5.1 Main features of BPN

3.5.1.1 The learning rate

A learning rate parameter selected for the network is of great effect on the net performance. This parameter must be of a small value to ensure the settlement of the net to a solution. A large value of learning rate may cause the net to skip the solution vector, and the too small value slows the learning process (Freeman and Skapura, 1991).

3.5.1.2 The momentum

The momentum factor is another feature that helps speeding up the net convergence. The term momentum stands for a fraction of the previous weight change added to the present one when calculating the weight change. The momentum acts to keep the direction of weight changes. The momentum factor ranges from 0 to 1. the larger the momentum the quicker the net steps towards convergence, so when large value of momentum is used, the learning rate should suitably be small to ensure that the net will not overshoot a solution (Freeman and Skapura, 1991) .

3.5.1.3 Bias and threshold terms

They are added in different problems and help in the convergence of the neural networks. They are dealt with as weights and their significance is realized by experience. If the bias term sign is negative, it is called threshold, though the sign makes no difference to the learning process. Anyhow, the use of bias terms is optional (Taylor, 1991).

3.5.2 Operation of BPN

BPN works by learning- in the first order- predefined examples of input- output pairs. The applied input pattern is propagated through the upper layers (a layer is a group of units or neurons at one level that are connected to the units in the previous and next layers) until an output is generated. The generated output is then compared to the desired one and the error signals are computed. These signals are transmitted back through the intermediate layer nodes and each unit in these layers receives a portion of the error signal that suits its contribution to the output layer. The point to be stressed here is that the intermediate layers arrange themselves to allow the recognition of the input features by their nodes which- when trained- will respond positively when the features of any input data resemble those of the input data during training (Freeman and Skapura, 1991).

So, BPN presents powerful means for the computer to examine incomplete data and recognize suitable patterns.

3.5.2.1 The delta rule

The generalized delta rule is the learning algorithm of network and it can be described mathematically.

Since the BPN is an interconnected- feed-forward layered network, there is no way for a feedback action to bypass any layer. The neural network that can compute a functional relationship between input and output is called a mapping network (Freeman and Skapura, 1991).

Consider Fig.(3.3), the net input to the j^{th} hidden unit is

$$net_{pj}^{h} = \sum_{j=1}^{N} w_{ji}^{h} x_{pi} + \theta_{j}^{h} \quad (3.2)$$

where;

w^h_{ji}=the weight in connection from the j^{th} input unit.

θ^h_j=the bias term.

x_{pi}= a vector component representing the output from the input layer.

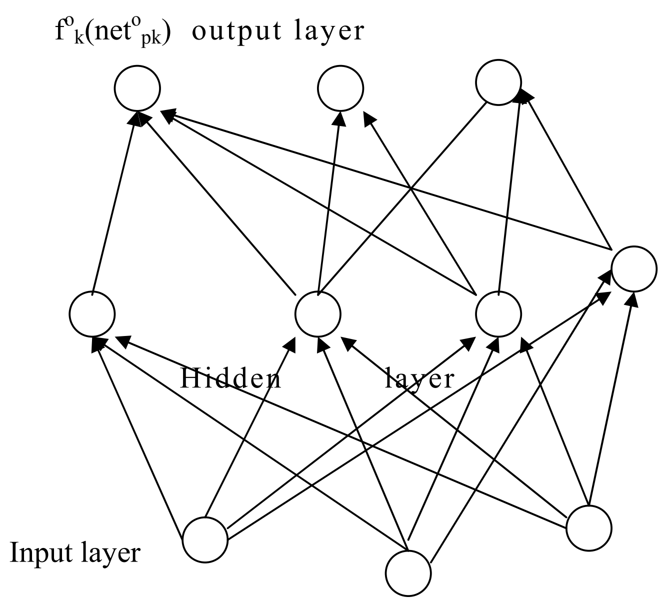


FIG. (3.3): network layers

Assuming the activation (input) of the node equals the net output yields,

$$i_{pj} = f^{h}(net^{h}_{pj}) \quad (3.3)$$

h refers to the hidden layer

The nodes output equations are

$$net^{o}_{pk} = \sum_{i=1}^{L} w^{o}_{ki} i_{pi} + \theta^{o}_{k} \quad (3.4)$$

$$o_{pk} = f^{o}_{k}(net^{o}_{pk}) \quad (3.5)$$

The superscript, o, refers to quantities of the output layer.

The output function $f^{o}_{k}(net^{o}_{pk})$ should be differentiable and give binary output units. In this case the sigmoidal function is used and it takes the following form:

$$f^{o}_{k}(net^{o}_{pk}) = \frac{1}{1+e^{-0.5 net^{o}_{pk}}} \quad (3.6\ )$$

Training the network can be summarized in the following steps:

1. Application of an input vector and calculation of the output values.
2. Comparison of the actual with the correct output and measurement of the error.
3. Determination of the change direction and amount of the weight for the purpose of reducing the error.
4. Application of the corrections to the weights and repetition of the above steps to all vectors and their errors reach an acceptable value.

3.5.2.2 Updates of the output-layer weights

The error at an output unit can be expressed as

$$\sigma_{pk} = (y_{pk} - o_{pk}) \qquad (3.7)$$

p is the p^{th} training vector and k is the k^{th} output unit

y_{pk} is the desired output value

o_{pk} is the actual output value

The error summarized by the generalized delta rule is

$$E_v = \sum_{k=1}^{M} \sigma_{pk}^{2} \qquad (3.8)$$

The direction of change of weights is obtained by calculating the negative gradient of E_p with respect to weights

∇E_p.

Then the weights values are readjusted such that the total error is reduced.

The difference between the delta rule and the least-squares technique is that the weights in the delta rule are changed with each training pattern and summed and a one update is made, while in the least-square technique the weight would not be changed until all the training pattern had been presented to the network.

3.5.2.3 Updates of hidden-layer weights

The same way of calculations that applied to the output layer can be followed, but the correct output of the hidden layer cannot be determined in advance, so measuring the output error is a difficulty.

The total error is expressed as

$$E_p = \tfrac{1}{2}\sum_k (y_{pk} - o_{pk})^2 = \tfrac{1}{2}\sum_k (y_{pk} - f_k^o(net_{pk}^o))^2$$

$$E_p = \tfrac{1}{2}\sum_k (y_{pk} - f_k^o(\sum w_{ki}^o i_{pj} \div \theta_k^o))^2 \qquad (3.9)$$

i_{pj} depends on the weights of the hidden layer.

So, the gradient of E_p can be written as:

$$\frac{\partial E_p}{\partial w_{ji}^h} = \nabla E_p = -\sum_k (y_{pk} - o_{pk}) \frac{\partial o_{pk}}{\partial (net_{pk}^o)} \frac{\partial (net_{pk}^o)}{\partial i_{pj}} \frac{\partial i_{pj}}{\partial (net_{pj}^h)} \frac{\partial (net_{pj}^h)}{\partial w_{ji}^h}$$

$$\nabla E_p = -\sum_k (y_{pk} - o_{pk}) \cdot f'^o_k(net_{pk}^o) w_{kj}^o \cdot f'^h_k(net_{pj}^h) x_{pi} \qquad (3.10)$$

Some practical considerations are to be noted in application of BPN:

a) Training the data

It is better to use as much as possible amount of data and preferably if they are noisy since this noise sometimes helps the network to converge.

The BPN switches off the similarity in the input data and ignores the irrelevant data, but the input data should cover the entire expected input data space since the BPN is inefficient in extrapolation.

b) Network sizing

There is no basis to set rules for choosing the number of layers or the units in one layer, but for hundreds or thousands of inputs a small size of hidden layers is needed. If the network fails to converge more nodes are needed in the hidden layers and if it does, the nodes can be reduced.

3.6 Applications of ANS (Keller, 2003)

3.6.1 The medical field

a) Medical diagnostic aides:

Artificial Neural Systems (ANS) proved to be good at diagnosing heart attacks. The significance of this fact is that, ANN processes large amount of data in emergencies. Some ANS technologies are employed in the diagnosis of cervical cancer by examination of pop smears.

b) Biochemical analysis

The ANS's have been used to analyse blood and urine, track glucose levels in diabetics, determine ion in body fluids and detect pathological conditions such as tuberculosis.

c) Medical image analysis

ANS use in the field includes classification of micro-calcifications in mammograms, classification of chest x-rays, tissue and vessel classification in magnetic resonance images, determination of skeletal age from x- ray images and determination of brain maturation.

d) Drug development

ANSs are used in developing drugs for treating cancer and AIDS as well as in the process of modeling biomolecules.

3.6.2 The nuclear science

a) Spectra analysis

ANSs are useful in analyzing lower resolution spectra. They are also, in the identification of radioactive isotopes which is very useful in automated identification of radioactive contaminants. They are also trusted to judge the spectra quality.

b) Modeling

ANSs are useful in modeling the nuclear stability, atomic mass and neutron energy.

c) Reactor controls and alarms

Some explorations were done in applying ANSs to reactor temperature control, core monitoring, alarm processing and identification of water level.

3.6.3 Environment applications

a) Identification of volatile chemicals

The ANSs are also applied as chemical sensing system and a pattern- recognition system to form an electronic “nose” that identifies the various volatile chemicals.

b) Environmental monitoring

Some explorations are running in the Pacific Northwest National Laboratory{6} to perform environmental restoration and waste management yielding to lower the cost of identifying the contaminants involved in toxic wastes, fuel mixtures, food industries …etc.

3.6.4 Energy applications

a) Diagnosis of turbine engines

Researches showed that the ANS can be employed effectively to monitor turbine engines performance data in real time and diagnose failure and faults in fuel system.

b) Operation and maintenance of central heating plants

An ANS based software is designed to increase the efficiency, reliability and safety of central heating plants at the US Marines Corps. bases. It helps the operators better control the plant configuration and operate within the design basis of the system.

Chapter Four

Data Collection and Application

Chapter (4)

Data Collection and Application

4.1 Data collection

The data, on which this thesis is based, is collected from the service reports of the Building and Road Researches Institute,(BRRI), University of Khartoum. For each combination of the parameters (LL, PI ..etc) the corresponding value of C_c is taken.

4.2 General classification of data

The total number of data samples used in constructing the network is 123 sets. The tests, from which these data resulted, are carried for different construction purposes (construction of buildings, road, bridges…etc). The data covered many areas in the Sudan as indicated in Table (4.1).

Table 4.1:Geographical distribution of the collected data

Area	No. of samples
Khartoum	82
West of the Sudan	21
The Gezira	2
Blue Nile	2
White Nile	11
East of the Sudan	3
River Nile State	2

The depths of these samples range from 1 to 25 meters. In choosing the data, no consideration was paid to sample being above, at or below the water table.

The soils, which these samples represent, are mixtures of sand, silt and clays with all their inter-classes whether based on Atterberg limits or sieve analysis results.

Also, there was no consideration to the load history of the samples. Some samples represent normally consolidated soils and some are over- consolidated.

This randomness of data- or noisiness- reflects the power of the neural network systems. On the other hands, it helps greatly in bringing the net to convergence.

4.3 Application and Methodology

The choice of the number of the layers is guesswork. The net consists of four layers: the input layer, the output layer and two hidden layers. The weights were initiated by random values (0.1) and the net readjusted them automatically. If the net undergoes a local minimum, the weights are revalued manually, until the absolute minimum is reached. The net is processed by using the Microsoft Excel solver. The developed network is named PRECUNNS; an abbreviation of Prediction of Compression Index Using Neural Network Systems)

The steps followed in constructing PRECUNNS are summarized below:

4.3.1 Parameters

The input parameters used are:

a) The liquid limit, LL

b) The plasticity index, PI

c) The percentage passing No. 200, %P

d) The initial moisture content, MC.

e) The bulk density, γ_b

f) The dry density, γ_d

g) The overburden pressure, σ_o

h) The void ratio, e_o

These parameters are normalized by the ordinary technique:

$$\text{Normalized value} = \frac{\text{actual value- minimum value}}{\text{maximum value- minimum value}}$$

4.3.2 Construction of the input to hidden layer

The input to any unit in the layer is termed net_i and calculated by the expression

$$net_i = \sum_{i=1}^{8} w_{ji} x_{pi} + \theta_j \quad ,\boldsymbol{j=1,8} \tag{4.1}$$

x_{pi}= the parameters; LL, PI, %P, MC, ……etc.

w_{ji}= the layers' weights.

net_i represents the contribution of each unit in the change of the weights.

4.3.3 Construction of the hidden layer

First, the input parameters, $\boldsymbol{i_{pi}}$,, are calculated using a sigmoid logistic function depending on the output of the previous layer (input layer)

$$i_{pi} = f_i^h(net^h) = \frac{1}{1+e^{-0.5net}} \quad \text{(equs. (3.3) and (3.6))} \qquad (4.2)$$

the new input for the next layer is

$$net_i = \sum_{i=1}^{8} w_{ji} i_{pi} + \theta_j \ \boldsymbol{,j=1,8}$$

and the new values of i_{pi} for the next layer are calculated by the same sigmoid function.

4.3.4 Construction of the output layer

The values of $\boldsymbol{net_i}$ (of the output layer) and the $\boldsymbol{o_p^o(f_i(net_i))}$ are calculated in the same manner as before.

The un-normalized values of the compression index are then calculated as follows:

$$C_{c-un-normalized} = C_{c-\min(actual)} + O_{pi}\left(C_{c-\text{maximum}(actual)} - C_{c-\min(actual)}\right)$$

4.3.5 Error computations

The error function is calculated as

$$Error\ function = (c_{c-NNW} - c_{c-actual})^2$$

The error percentage is computed by the expression

$$Error\% = \frac{(C_{c-NNW} - C_{c-actual}) \times 100}{C_{c-actual}}$$

The subscript NNW refers to the values obtained from the neural network

The preparation ends by computing the maximum, minimum, Chi- test and the average of percentage error.

Then the software solver is used after setting the initial weights at 0.1 (random values) and the sum of the error function is minimized to the smallest possible value.

Chapter Five

Results and Analysis

Chapter (5)

Results and Analysis

5.1 Introduction

It is understood that the results obtained by mathematical means shall not always be an identical image of the actual values. The approximations in the later and the natural erroneous factors involved in the former made a good job in preventing the occurrence of such ideal case. So, the speech is about how much the theoretical results come near the actual ones.

In neural networks it is impossible to predict this nearness. The process aims at the minimum possible error. The identification of this minimal error is established by observing the marching of the net towards convergence.

In this thesis, for more confirmation, a regression analysis is applied to the data to be compared to PRECUNNS results and both will be displayed in this chapter followed by a parametric analysis of the results.

5.2 Regression analysis results

A series of correlations between the compression index and its more affecting parameters were carried out tracking the previous works explained in Chapter (2). These correlations are carried by using the software SPSS11.0 for Windows. The results of these correlations are shown in Table (5.1)

Table (5.1): Summary of the correlations between c_c and parameters (obtained by using the software SPSS 11.0)

c_c versus	R^2 value
LL	0.01
e_o	0.53
MC	0.351
e_o and MC	0.272
PI and e_o	0.53
LL and e_o	0.53
PL and γ_d	0.475
LL and γ_d	0.474
LL, PI and γ_d	0.023

Of all these correlations only the relations between the compression index and e_o, the combination of e_o and PI, and the combination of e_o and LL resulted in values of R^2 greater than 0.5(Figures (5.1)-(5.3)).The expressions of these relations are:

$$c_c = 0.45\ e_o - 0.111 \tag{5.1}$$

$$c_c = 0.452\ e_o - 0.000351\ PI - 0.122 \tag{5.2}$$

$$c_c = 0.451\ e_o - 0.0001459\ LL - 0.104 \tag{5.3}$$

5.3 PRECUNNS results

5.3.1 Layers components

The input layer consisted of 8 neurons (LL, PI,...,e_o). Their values are normalized and then applied to the net. The two hidden layers consisted of 8 neurons each. And the output layer consisted of two neurons; one is the compression index and the other is recompression index.

The activations between the layers progress through a set of 9 weights (one of them is the bias term).

5.3.2 Compression index values

The net was trained with 123 sets of data from experimental work. The training process had yielded a solution, which is shown in Table (5.2). The values of the compression index from the net are compared to those from the experiment and plotted in Figure (5.4)

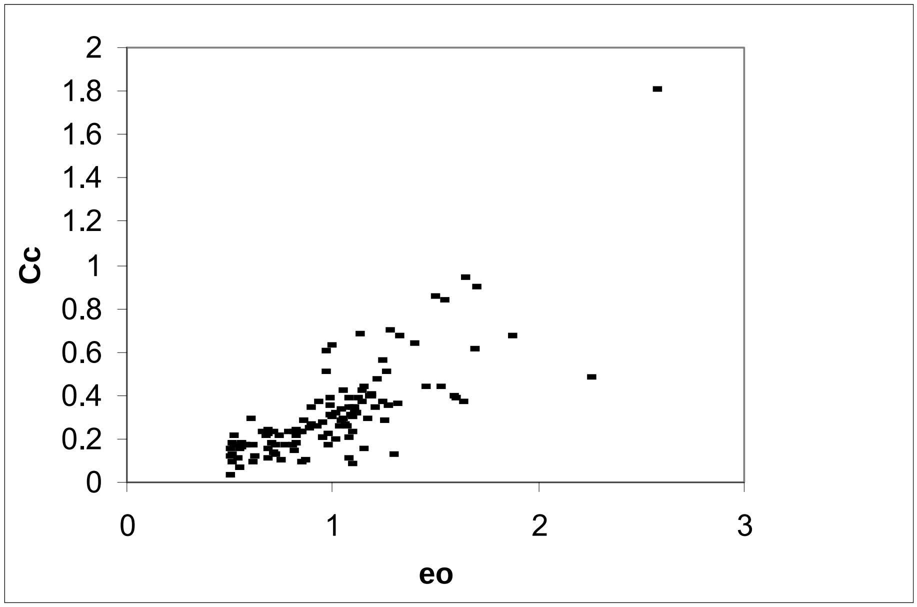


Fig (5.1): Relationship between compression index, C_c and void ratio, e_o

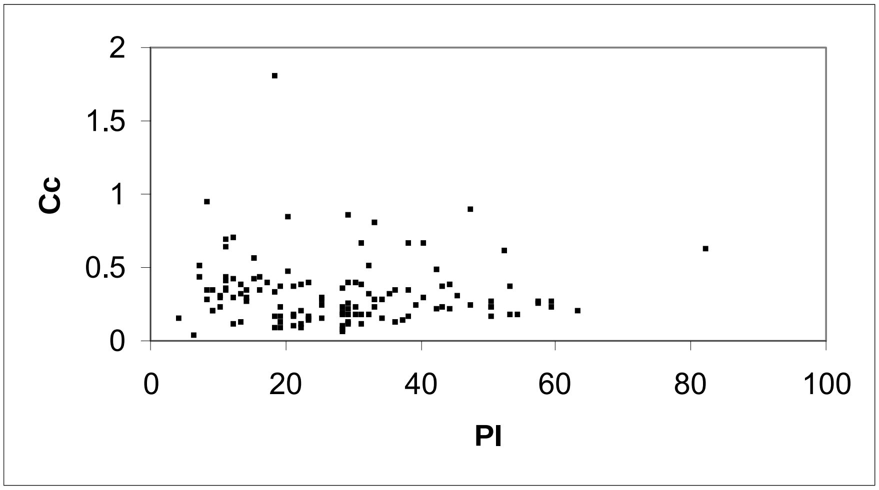


Fig.(5.2): Relationship between C_c and PI

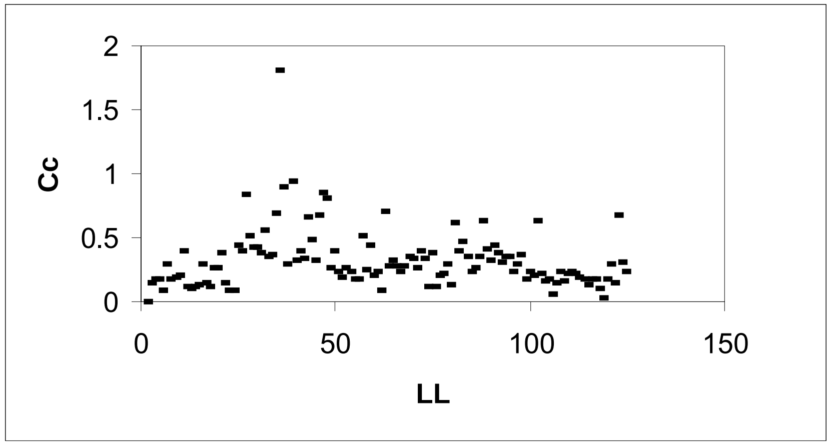


Fig.(5.3): Relationship between C_c and LL

Table (5.2):Experimental and PRECUNNS values of C_c

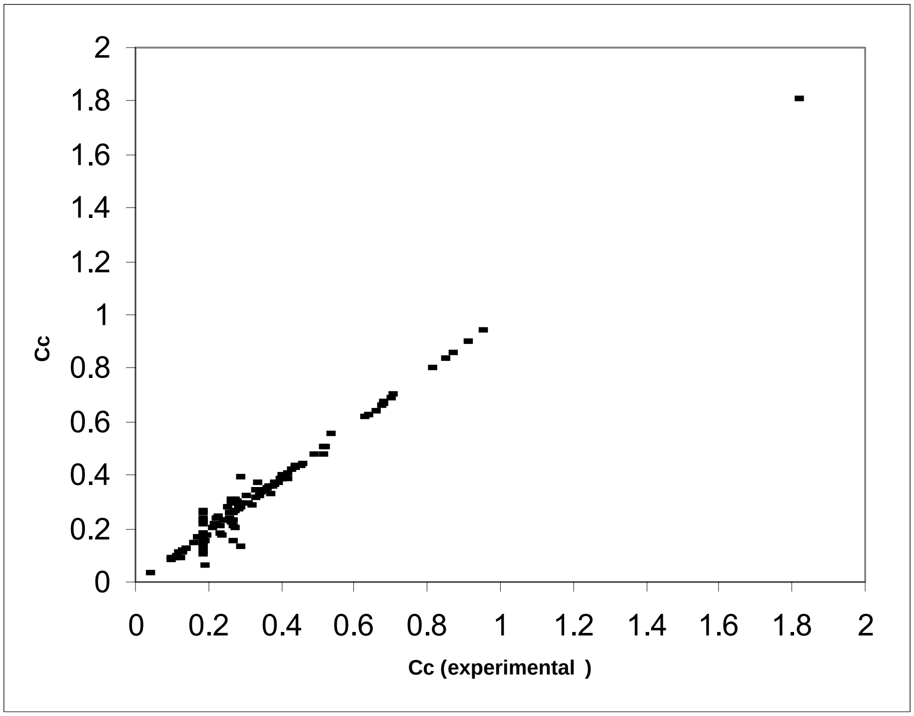
No.	C_c(NNW)	C_c(Exp.)	No.	C_c(NNW)	C_c(Exp.)
1	0.25891	0.1528	44	0.6726	0.67255
2	0.175564	0.17	45	0.859919	0.86
3	0.229149	0.1733	46	0.802238	0.803
4	0.088594	0.08276	47	0.260135	0.26
5	0.254757	0.2989	48	0.39396	0.39
6	0.161942	0.172	49	0.244904	0.23
7	0.22314	0.1856	50	0.175581	0.186
8	0.202801	0.2071	51	0.245518	0.26
9	0.388443	0.3905	52	0.175566	0.23
10	0.119747	0.1147	53	0.175633	0.175

11	0.175742	0.1053	54	0.17554	0.17
12	0.175637	0.1109	55	0.503649	0.509
13	0.175589	0.1336	56	0.214391	0.249
14	0.271358	0.293	57	0.429957	0.438
15	0.175554	0.1412	58	0.220661	0.211
16	0.175569	0.123	59	0.255331	0.232
17	0.17555	0.268	60	0.111409	0.091
18	0.177271	0.2638	61	0.696204	0.701
19	0.389858	0.389	62	0.270818	0.286
20	0.148239	0.1468	63	0.331575	0.322
21	0.099315	0.093	64	0.280318	0.281
22	0.088826	0.093	65	0.230581	0.231
23	0.442708	0.44	66	0.242379	0.283
24	0.39074	0.402	67	0.357472	0.358
25	0.840155	0.84	68	0.343724	0.344
26	0.509479	0.508	69	0.271023	0.272
27	0.419733	0.422	70	0.391532	0.392
28	0.42591	0.428	71	0.317237	0.342
29	0.407898	0.387	72	0.108871	0.114
30	0.523241	0.559	73	0.379692	0.376
31	0.343118	0.346	74	0.117202	0.117
32	0.370241	0.369	75	0.264437	0.203
33	0.688315	0.687	76	0.175553	0.22
34	1.807	1.807	77	0.311753	0.287
35	0.900718	0.9015	78	0.279743	0.132
36	0.297739	0.2979	79	0.617992	0.618
37	0.942474	0.9427	80	0.279475	0.396
38	0.36288	0.3286	81	0.504915	0.476
39	0.402812	0.4	82	0.3642	0.36
40	0.332968	0.34	83	0.253806	0.228
41	0.663785	0.664	84	0.264573	0.265
42	0.481209	0.481168	85	0.3409	0.348
43	0.291924	0.3245	86	0.647221	0.638

Table (5.2) continued

No.	C_c(NNW)	C_c(Exp.)	No.	C_c(NNW)	C_c(Exp.)
87	0.40494	0.405	106	0.229697	0.231
88	0.319326	0.317	107	0.173506	0.157
89	0.44678	0.441	108	0.212703	0.22
90	0.370215	0.376	109	0.210445	0.242
91	0.260934	0.311	110	0.205875	0.216
92	0.34874	0.35	111	0.175562	0.185
93	0.350751	0.352	112	0.170366	0.17
94	0.245406	0.236	113	0.130104	0.13
95	0.266282	0.3	114	0.17562	0.174
96	0.322901	0.37	115	0.175565	0.175

97	0.183394	0.1772	116	0.100536	0.1
98	0.175575	0.2425	117	0.033002	0.033
99	0.260208	0.208	118	0.175573	0.17
100	0.628016	0.628	119	0.276646	0.296
101	0.175564	0.226	120	0.175807	0.153
102	0.171991	0.166	121	0.671124	0.671
103	0.175605	0.1825	122	0.254081	0.312
104	0.179483	0.065	123	0.175541	0.234
105	0.180713	0.154			



Figure(5.4): C_c from PRECUNNS versus C_c from experiments

5.3.3 Network weights:

The initial guess value of the weights for the output and both hidden layers was 0.1 all through. These values had been changed by the net to settle at the final values shown in Tables (5.3), (5.4) and (5.5) for C_c and Tables (5.6), (5.7) and (5.8) for C_r.

Table (5.3): Values of weights of the first hidden layer (for C_c)

w_1	w_2	w_3	w_4	w_5
-27.2638	-37.5591	-2.32484	-47.5554	41.2104
8.196932	-19.9626	4.48865	-57.1309	7.113211
1.019252	-19.6416	28.39077	-29.3578	-3.53177
0.907389	6.804896	14.10614	3.756648	26.51712
65.36065	-6.51683	36.94621	-13.8257	-20.9355

-14.8354	-26.097	19.84486	-15.5654	-30.2851
28.57877	8.535233	8.253802	-3.05894	-6.52296
5.272007	-35.118	-3.44586	-10.7515	-13.7707

w_6	w_7	w_8	θ
39.62999	-4.18087	17.23982	21.86259
-0.26388	49.30593	-14.8111	12.20858
19.75845	47.93452	-13.4052	4.986686
9.522165	13.59291	-6.66759	6.678984
-0.38886	-15.992	-18.4913	5.072184
25.88656	24.66232	22.26055	-12.3122
2.032957	-5.99449	-12.0105	7.534791
-8.67229	12.59826	6.658979	-12.0883

W_1, W_2,...,W_8 are the final values of weights

θ is the bias term

Table (5.4): Values of weights of the second hidden layer (C_c)

w_1	w_2	w_3	w_4	w_5
43.79611	7.323221	4.74277	8.673085	-6.83208
-12.2344	51.96866	-17.8228	-18.7624	31.17131
-0.69317	-4.2543	18.87284	4.29235	-3.38203
16.91284	-84.3389	31.18786	14.94899	-50.1988
3.222877	3.986021	3.68302	9.958683	0.448254

16.81975	46.10631	5.579323	-0.38262	6.383373
21.90711	25.64051	18.74385	-43.3917	40.72461
4.963826	-15.7087	-6.01522	8.289831	-29.1315

w_6	w_7	w_8	θ
-45.9636	-18.3121	16.39108	13.55898
-66.7402	-13.6258	12.83269	3.475368
18.84459	0.598941	-19.0117	10.31483
118.8218	15.59747	-18.5935	15.1706
27.07503	-0.84925	-1.36411	12.18627
38.82976	-24.8311	-9.363	1.158527
6.427863	-46.3513	0.962703	10.36021
-9.47534	8.950229	-1.61505	9.851724

W_1, W_2,…,W_8 are the final values of weights

θ is the bias term

Table (5.5): Values of weights of the output layer (for C_c)

W_1	w_2	w_3	w_4	w_5
-8.50368	-58.9366	12.60456	-57.9552	19.37904

0.1	0.1	0.1	0.1	0.1

w_6	w_7	w_8	θ
29.45911	-28.1244	-29.1705	29.24907
0.1	0.1	0.1	0.1

W_1, W_2,…,W_8 are the final values of weights

θ is the bias term

Table (5.6):Values of weights of the first hidden layer (for C_r)

w_1	w_2	w_3	w_4	w_5
-33.4286	-44.1395	19.16722	-73.5847	112.6096
-86.2541	-59.6056	35.7001	-9.78684	-99.9547
3.516739	-15.6483	33.45607	-22.1358	-1.2121
0.91092	6.804544	14.11881	3.748812	26.51829
4.869747	-35.2128	-5.29924	-11.6222	-14.1325
-54.3688	52.6938	10.5891	-112.825	22.37275
15.68182	34.84298	21.61127	-21.1237	5.684478
67.72492	-5.56605	51.26691	-16.2191	-24.5858

w_6	w_7	w_8	θ
47.20768	-19.7931	-23.4273	18.55089
94.51935	16.12508	-7.94225	23.5405
22.02287	47.88035	-6.67139	14.47088
9.526139	13.59828	-6.6728	6.681492
-9.32335	12.01293	5.193961	-14.6467
-57.5167	78.07304	-144.86	65.987
-29.3391	15.75662	-12.2377	6.664713
-1.3947	-25.7761	-14.4736	13.71605

W_1, W_2,…,W_8 are the final values of weights

θ is the bias term

Table (5.7): Values of weights of the second hidden layer (C_r)

w_1	w_2	w_3	w_4	w_5
44.21164	7.250855	5.349165	8.368817	-6.86819
-15.456	103.171	-33.9209	-32.66	29.99098
0.54734	-2.84444	20.09258	6.452226	-3.39341
21.14628	-57.7269	25.60999	16.10776	-47.5054
3.186121	3.983573	3.648875	9.92455	0.446039
11.47166	82.70209	7.185037	2.761844	6.841694
21.90608	33.70803	15.27036	-37.9935	38.83061
27.26705	-16.2965	-8.55074	4.542693	-27.6147

w_6	w_7	w_8	θ

-46.2156	-15.0891	16.53764	13.37363
-94.8435	7.313055	-7.94807	-16.8736
18.89284	2.697347	-16.8706	12.47472
110.445	-3.25846	-19.5876	16.30414
27.07479	-0.88282	-1.39596	12.15212
37.19134	-19.7127	-4.25129	3.964623
-15.8097	-67.7655	18.23898	2.722265
-35.9007	23.83098	-0.27123	6.557786

W_1, W_2,…,W_8 are the final values of weights

θ is the bias term

Table (5.8): Values of weights of the output layer for C_r

w_1	w_2	w_3	w_4	w_5
-9.58211	-54.5011	10.3939	-53.5068	18.34735
6.262762	-21.296	6.209444	-26.4956	6.236751

w_6	w_7	w_8	θ
28.47089	-27.9482	-29.08	29.93708
-4.97833	-3.56405	1.562066	6.219481

W_1, W_2,…,W_8 are the final values of weights

θ is the bias term

5.4 Test of PRECUNNS

The net is tested to verify its output. Data of 16 sets is used (Table 5.9). The neural network values are shown together with experimental in Table 5.10.

Table (5.9):Test data

	LL	PI	%P	MC	γ_b	γ_d	σ_o	eo	Cc
1	46	19	88	40.89	18.25	12.95	107	1.12	0.48
2	40	14	77.1	33.11	19.77	14.88	170	0.85	0.23
3	61	28	34	90	29.2	15.67	12.13	47	1.235
4	43	19	13	90.7	20.5	14.19	11.78	110	1.284
5	82	23	52	97.8	56.3	15.81	10.12	72	1.678
6	47	34	17	88.1	19.3	14.78	12.39	120	1.163
7	50	39	20	97.7	44.1	17.62	12.23	80	1.2
8	42	26	11	82	38	16.12	11.68	135	1.295
9	76	31	47	70.7	25.2	17	13.5	102	1
10	74	31	51	52.8	19	20.4	17.1	98	0.549
11	47	27	26	63.5	20.1	17.8	14.8	59	0.786
12	62	26	31	74.8	35.5	18.9	14	57	0.935

13	66	26	36	55	33.3	19.1	14.3	86	0.884
14	58	30	30	60	27.2	18.7	14.7	56	0.839
15	54	20	29	65	27.3	19	15	114	0.811
16	74	27	40	70	25.5	18.1	14.4	82	0.875

Table (5.10): C_c values obtained by PRECUNNS and the test data values.

	$c_{c(NNW)}$	$c_{c(EXP)}$
1	0.474017	0.48
2	0.254709	0.23
3	0.450346	1.235
4	0.279743	1.284
5	0.617992	1.678
6	0.279475	1.163
7	0.504915	1.2
8	0.3642	1.295
9	0.177531	1
10	0.175555	0.549
11	0.175627	0.786
12	0.255345	0.935

13	0.279366	0.884
14	0.21347	0.839
15	0.175558	0.811
16	0.190975	0.875

5.5 Comparison between the regression analysis and neural networks

On the basis of the work carried in this thesis, comparing between the regression analysis and the neural network technique resulted in the following:

1- The approximations in the regression analysis are less accurate than in the neural network.
2- The solutions in the regression analysis are obtained with lesser time and smaller effort.
3- In regression analysis, the misread data lead to unreasonable solution (see Figs (5.1) to (5.3)), whereas these values help the neural network to reach the solution.

Chapter Six

Summary and Conclusion

Chapter (6)

Summary and Conclusion

6.1 Summary

A considerable amount of the experimental data involved in determination of the compression index of different soils was collected. The data used in this study is collected from the service reports of the Building and Road Researches Institute, University of Khartoum. The sources of the data cover most of the Sudan. The data is normalized and then the net is constructed.

This thesis used the neural network technique to evaluate the compression index. Linear regression based on least square technique was also used. With this skeleton, the data collected is used to "train" the net so that it becomes able to evaluate the compression index for any set of data. The most common type of neural network is the back- propagation net which uses the supervised training, was used.

6.2 Conclusion

Some beneficial conclusions can be inferred considering both the neural network and regression results.

6.2.1The neural network results

Although the data used for the purpose of constructing the neural network is somewhat raw, the results obtained are reliable. The error of the resulting values of the compression index is acceptable(see Table(5.2)). The model was validated using laboratory data and the outcome is acceptable.

6.2.2 The regression analysis results

1. The results of the thesis showed that the compression index is affected by the liquid limit and the relation is linear.
2. The void ratio affects the value of the compression index, the relation is a direct proportionality.
3. The moisture content affects the compression index in the same manner as the void ratio.

4. An investigation of the relation between the void ratio, the moisture content and the compression index confirmed the existence of a linear relationship between the ratio $C_c/(1+e_o)$ and the moisture content.

   The best fit was found between C_c and e_o, C_c and the combination of LL and e_o and C_c and the combination of PI and e_o. These correlations gave a maximum value of R^2=0.53.

6.3 Recommendations

The work done in the thesis cannot be claimed to be complete. Some recommendations are attached here that are expected to promote the results obtained in this thesis, and they are:

1. A further training process can be tried to obtain better results, by some marginal changes in the net weights
2. The net can be reconstructed by using refined data; by classifying the soil types. The regression analysis can also be tried.
3. Also the net may be reconstructed and tried by using only

the parameters that affect the compression index mostly.

4. the idea of thesis can be applied to predict other soil properties such as the shear parameters c (cohesion of the soil) and ϕ (the angle of shearing resistance).

References

1- Craig, R. F., 1978, “Soil Mechanics”, 4th edition, Van Nortran Renold, England.

2- Freeman, J. A., and. Skapura, D. M., 1991,“ Neural Networks: Algorithms, Applications and Programming Techniques” , First edition, Addison-Wesley Publishing Company, USA.

3- Herrero, O. R, 1980, “Journal of the Geotechnical Engineering”, “Universal Index Equation”, Vol.106, Nov.1980.

4- Holtz, R. D., and Kovacs, W. D., 1981,” An Introduction to Geotechnical Engineering”, 1st edition, Prentice- Hall, Englewood Cliffs, New Jersey.

5- Keller, P. E. The Pacific Northwest National Lab, www.emsl.pnl.gov.

6- Taylor, J. G., 1995, “Neural Networks”, 1st edition, Alfred Waller Ltd. and Unicom Ltd.

7- Terzagi, K., Peck, R. B. and Mesri, G., 1996, “Soil Mechanics in Engineering Practice”, 1st edition, John Wiley and Sons.

.